BEBR

FACULTY WORKING
PAPER NO. 888

The Impacts of Skewness and Kurtosis on the Risk Estimation and Determination.

Cheng F. Lee
Chunchi Wu

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois, Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 888

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

August 1982

The Impacts of Skewness and Kurtosis on the Risk Estimation and Dètermination

Cheng F. Lee, Professor
Department of Finance

Chunchi Wu
Syracuse University

Abstract

Based upon modern statistical theory and econometric methods, this study investigates the impacts of skewness and kurtosis on risk estimation and determination. It was found that both skewness and kurtosis of security rates of return are important in capital asset pricing determination.

I. Introduction

It has been known for some time that the returns of security and the residuals from market model regressions are not exactly normally distributed. Mandelbrot (1963) first demonstrated that stock return did not follow the normal distribution. Mandelbrot (1963, 67) and Fielitz (1971) showed that the variation in stock prices is not stationary. Fama (1965, 76) and Mandelbrot suggested a stable class distribution as the probability model for the distribution of stock prices. Fama estimated the characteristic exponents for the stable Paretian distribution to be somewhere from 1.7 to 1.9 in contrast to a value of 2.0 for the normal distribution and 1.0 for the Cauchy distribution. Blume (1968) also provided empirical evidence that a characteristic exponent between 1.7 and 1.8 summarized the residuals from market model regressions rather well.

On the other side, Officer (1972) found that the standard deviation of the distribution of daily and monthly stock returns is a well behaved dispersion measure. This finding is inconsistent with the findings of stable Paretian distribution by Fama and others, since the standard deviation is expected to be large and behave erratically if the stock prices follow a stable class distribution. Officer also provided evidence that some of the properties of the stock return distribution are not consistent with the stable hypothesis. Specifically, the time-series sums of daily stock returns became "thinner-tailed" for larger sums, and the cross-sectional sums of monthly stock residuals from the market model had a characteristic exponent larger than that of

component stocks. Hsu, Miller and Wichern (1974) chose the minimum chi-square goodness-of-fit procedure to test the stable hypothesis. Their results for four companies showed that symmetric stable distributions are not consistent with the daily data although the stable Paretian model fits better on the monthly rate of returns. Based upon their findings, Hsu, Miller and Wichern (74) proposed a normal probability model with a nonstationary variance subject to step changes at irregular time points corresponding to shifts of various exogenous factors.

The determination of stock return distribution is important for the risk estimation. Fisher (1932) showed that the dispersion of sample variance is a function of the fourth moment. This implies the measured total risk of stocks will be affected by the value of kurtosis which appear to be positive much often than negative.

The nonstationarity of stock returns probability distributions and distribution parameters have led to the instability of beta systematic risk estimates. Francis (1979) showed that the stability of beta coefficient estimated for each individual security can be affected by three components: (1) correlation coefficient between the individual security and the market portfolio; (2) the standard deviation of individual security returns; and (3) the standard deviation of market returns. In a three parameter capital asset pricing model, Kraus and Litzenberger (1976) and Lee (1976) showed that both beta systematic risk and gamma systematic skewness estimated from the quadratic characteristic lines are partly determined by variance, skewness, kurtosis and the investment horizon of the market returns. Therefore, to resolve

the problem of risk estimation and understand the short run and long run stability of risk measures, we need to closely examine the behavior of skewness and kurtosis.

In the estimation of beta systematic risk, the least squares estimator will be best linearly unbiased only if classical assumptions on linear model are satisfied. One of these assumptions is the normality of the observations. If this assumption is violated due to, for example, the existence of skewness and high kurtosis, then the least regression (LS) can no longer provide the best linear unbiased estimator. Under this condition, other robust estimators may be superior to the LS estimator, unless either the original data are transformed to satisfy the classical assumptions of the linear model, or the skewness and kurtosis are not serious.

It is well-known that the mean-absolute-deviation (MAD) method has been suggested as an alternative to the least squares regression in estimating the regression parameters for a fat-tailed distribution. The appeal of MAD apparently comes from the fact that it gives relatively less weights to outliers as compared to LS regression. Theoretically the MAD estimator should be more efficient than LS estimator when applied to the fat-tailed distributions. However, the empirical reports on the performance of MAD estimator varied widely. Meyer and Glauber (1964), Winginton (1968) found that MAD regression outperformed or at least provided no worse forecasting power than the LS regression. Ruppert and Carroll (1980) reported that MAD will be more efficient only if the kurtosis is very high. On the other hand, Wise (1963) argued that neither LS nor MAD can provide the best linear unbiased

estimator for a characteristic exponent with values between 1 and 2. Sargent (1969) used simulation to show that for distributions with characteristic exponents greater than 1.7, LS regression performs better than MAD regression. Two recent studies which are directly related to the systematic risk estimation showed no sign of improvement from MAD regression. Sharpe (1971) provided evidence that two methods give similar results. Cornell and Dietrich (1978) found that the short term instability or beta estimates does not resulted from the fat-tailed distribution. Their study further supported Sharpe's and Sargent's findings.

However, these comparisons should be taken more cautiously. When the skewness of individual security return is significant, the efficiency of MAD regression should be reduced despite the existence of high kurtosis. Furthermore, the sequential parameter variation may be due to various problems such as structural changes, misspecifications, aggregation, and error in variable. If these probelms are serious, the performance of MAD method become difficult to evaluate before problems are removed.

The purpose of this study is to examine the impacts of skewness and kurtosis on the determination and stability of risk measures. The classical two parameter model and the three parameter models recently proposed are employed to estimate different risk measures. Section 2 discusses the stability of total risk. The theoretical relationship between total risk and the higher moment is laid out and the intertemporal stability of total risk is examined. Section 3 examines the stability of beta risk for each subgroup with similar

return distributions. Section 4 explores the impacts of skewness and kurtosis on risk decomposition. Possible impacts of skewness and kurtosis on the estimated betas are also investigated in accordance with quadratic market model. To complete the investigation on the effect of distribution parameters on the risk estimates, Section 5 evaluates the performance of mean-absolute-deviation and least squares estimators when the rates of return are not normally distributed. Finally, the results of this study are summarized in Section 6.

II. The Stability of Total Risk

Recent research in risk measures has indicated that the variance of security return is more stable than the beta systematic risk. Francis (1979) showed that most of the standard deviations of individual security returns are fairly stable intertemporally, and suggested that the historical standard deviation may be used without adjustments for intertemporal change in the valuation of put and call options. The relative stability of standard deviation also implies that Sharpe's portfolio performance measure may be superior to Treynor's or Jensen's.

The stability of risk statistics is an empirical question. Nevertheless, theoretically the higher moment of returns distributions can affect the sample variance. If security returns are not normally distributed, the estimated sample variance of security returns is more unstable over time than that of the normal distribution. Fisher (1932) showed that for any infinite population, the variance of sample variance from random samples of size n can be written as

$$V(s^2) = \frac{2\sigma^4}{n - 1} \left(1 + \frac{n - 1}{2n} \frac{k_4}{\sigma^4}\right) \qquad (1)$$

The first component is the variance of sample variance given that the parent distribution is normal. The second component represents the influence of non-normality. k_4 is Fisher's fourth cumulant defined as

$$k_4 = E(Y_i - E(Y_i))^4 - 3\sigma^4 \qquad (2)$$

This fourth cumulant is zero for a normal distribution. Notice that the stability of sample variance is only affected by the relative kurtosis, and is not affected by skewness. Also the effect of relative kurtosis is almost independent of the sample size and will remain even with large sample. Therefore, using only the sample variance as the major criterion for the stability of risk measures may be misleading if kurtosis of the distribution is not considered.

The stability of total risk is reexamined in this section. Monthly returns for 464 securities and the market portfolio are collected from the CRSP tape. The sample period is from January 1959 to December 1979. Data are first grouped into three nonoverlapping periods: 1959-65, 1966-72, 1973-79, yielding 84 return observations for each period. The monthly treasure bill rate is used as a proxy for the risk free rate. Company data are further divided into two groups according to the degree of kurtosis of returns distribution. The regression model is specified as

$$s^2_{j,t+1} = a_0 + a_1 s^2_{j,t} + u_{t+1} \qquad t = 1,2 \qquad (3)$$

The sample variance of security in the current period is regressed against the sample variance in the last period. Equation (3) is estimated for three groups: normality, significant kurtosis, and all (464) companies. Mincer and Zarnowitz's (1969) mean square error (MSE) method is employed to decompose the sources of error. Results are displayed in Table 1.

The effect of kurtosis on the stability of total risk is significant. The values of R squares drop substantially for the group with significant kurtosis when the sample variance in the previous period is used to estimate the variance in the current period. The total mean square errors for the group with significant kurtosis are much larger than those for the group with normality. In the first period regression, the total mean square error for the second group is about 2.2 times the size of the first group. In the second period, this ratio goes up to 2.8. Overall, the major estimation errors are due either to random variation or bias, while the inefficiency term is generally small. The inefficiency term is slightly larger for the second group with significant kurtosis. Essentially, inefficiency measures the deviation of slope value from one. As it can be seen, the slope coefficients of the second group regression are significant smaller than one. Thus, the results in Table 1 suggests that the kurtosis of returns distribution must be considered when the historical standard deviation is used to estimate the total risk. These results are consistent with Cootner's (1962) findings that relative kurtosis information is important for security analysis. They also indirectly support Scott and Horvath (1980)

Table 1

Summary Statistics for the Stability of Total Risk

Group	Slope	R^2	Mean Square Errors*			
			Bias	Inefficiency	Random	Total
(1) Normality						
1959-65 1966-72	1.167	.53	.0000058 38%	.0000002 1%	.0000093 61%	.0000153
1966-72 1973-79	.952	.56	.0000038 21%	.0000000 0%	.0000142 79%	.0000181
(2) Significant kurtosis						
1959-65 1966-72	.781	.39	.0000045 13%	.0000014 4%	.0000281 83%	.0000339
1966-72 1973-79	.846	.45	.0000063 13%	.0000010 2%	.0000397 85%	.0000470
(3) All Companies						
1959-65 1966-72	.917	.45	.0000054 26%	.0000001 0%	.0000156 74%	.0000211
1966-72 1973-79	.921	.53	.0000044 17%	.0000001 0%	.0000210 83%	.0000255

*The components of mean square error are reported in both absolute and percentage terms.

theoretical analysis on the direction of preference for moments of higher order than the variance.[1]

III. The Stability of Beta Risk

The beta risk is determined by three portions: standard deviation of individual security and market portfolio returns, and correlation coefficient between individual security and market portfolio returns. Hence, the variance (stability) of beta is affected by the variances of these three components. Beta therefore is intrinsically more unstable than the total risk of individual securities. To investigate the stability of beta, equation (4) is used as the estimate model. Again, regressions are estimated for three groups: normality, significant kurtosis and all companies.

$$\text{(4)} \qquad \text{Beta}_{j,t+1} = b_0 + b_1 \text{ Beta}_{j,t} + v_{t+1} \qquad t=1,2$$

Table 2 shows the regression results. The slope of equation (4) is significant less than one in every case. The values of R squares and slope coefficients are relatively small as compared to the results of total risk regressions. The kurtosis affects the stability of beta risk indirectly through its effect on the total risk. The results for the second group with significant kurtosis is consistent with this line of reasoning. The R squares and estimated slopes are smaller, and the total mean square errors are larger for the second group than for the first group.

Note that the impact of kurtosis on the total risk is more significant than on the systematic risk. Another interesting result can be found in the structure of mean square errors. Unlike total risk

Table 2

Summary Statistics for the Stability of Systematic Risk

Group	Slope	R^2	Mean Square Errors			
			Bias	Inefficiency	Random	Total
(1) Normality						
1959-651 1966-72	.608	.29	.0094926 7%	.0190702 13%	.1119369 80%	.1404997
1966-721 1973-79	.487	.35	.0011628 1%	.0402199 37%	.0665661 62%	.1079488
(2) Significant kurtosis						
1959-651 1966-72	.598	.24	.0091584 5%	.0225179 12%	.1539511 83%	.1856275
1966-721 1973-79	.495	.33	.0046785 3%	.0574870 33%	.1118450 64%	.1740105
(3) All Companies						
1959-651 1966-721	.605	.27	.0093896 6%	.0201048 13%	.1251045 81%	.1545990
1966-721 1973-79	.489	.34	.0018404 1%	.0448647 36%	.0781831 63%	.1248882

regression, the inefficiency term becomes a major source of forecasting errors. The relatively high inefficiency implies that beta estimates are intertemporally more unstable than the variance of returns.

IV. The Impacts of Skewness and Kurtosis on Risk Decomposition

In this section, the estimate model consistent with the traditional two parameter capital asset pricing is introduced, and the problem of functional specification is investigated. Three versions of Box-Cox transformation are used to remove the skewness and kurtosis of the return probability distribution. Following this, the impacts of skewness and kurtosis on the specification of the model and the distribution of disturbance are examined more closely. Finally, the model developed is linked to the three parameter capital asset pricing model.

1. Generalized Functional Form

The linear model based on the traditional capital asset pricing theory may not be an appropriate functional form to explain the risk return relationship if the returns of securities are not normally distributed. Nevertheless, there may exist a transformation such that the transformed observations are normally distributed. Two methods are considered in this section. The first method is the Box-Cox transformation, and the second is the CES function approximation.

A. Box-Cox Transformation

A special class of transformation developed by Box and Cox (1964) is useful for inducing normality on observations from skewed distributions.

In empirical econometrics, the Box-Cox transformation has been primarily used as a device for generalizing functional form. Following Fabozzi, Francis and Lee (1980) [FFL], a generalized functional form of specifying the market model can be written as

$$R_{it}^{\lambda} = a + b\, R_{mt}^{\lambda} + e_{it} \tag{5}$$

where λ's are the functional form parameters for the dependent and independent variables. Equation (5) reduces to the linear form when λ is equal to one; and becomes the log-linear form when λ approaches zero. Hence equation (5) includes both linear and log-linear forms as special cases and provides a generalized functional form to test whether the linear model is appropriate for investigating the risk-return relation. There are at least two interpretations for λ. Statistically, in a narrow sense lambda is the functional transformation parameter used to eliminate the skewness and kurtosis of the dependent variable, for the present case, the rates of returns of individual securities. In a broader sense, the value of λ determines the correct specification of the regression model which can either be linear or nonlinear.

In the finance literature, λ measures the ratio of the true investment horizon to the observed horizon in the equilibrium risk return relationship as used by Jensen (1969) and Lee (1976). When the value of λ approaches zero, the market equilibrium is instantaneous, provided that the returns are defined over infinitesimally small time intervals. Jensen (1969) has developed instantaneous systematic risk by setting λ equal to zero.

To perform the transformation, FFL suggests that equation (5) can be expressed as

$$(6) \qquad r_{it} = a' + br_{mt} + \varepsilon_{it}$$

where

$$r_{it} = \frac{R_{it}^{\lambda} - 1}{\lambda}$$

$$r_{mt} = \frac{R_{mt}^{\lambda} - 1}{\lambda}$$

$$a' = \frac{(a+b) - 1}{\lambda}$$

$$\varepsilon_{it} \sim N(0,\sigma_{\varepsilon}^{2})$$

The residual analysis method considered by Box and Cox is used to estimate a parameter which primarily measures the degree and the direction of skewness. The value of this parameter for each individual security is approximately equal to equation (7).

$$(7) \qquad \alpha = \frac{M_1M_3 - \frac{1}{3}\left[(M_4 - 3M_2^2) + \frac{3M_2M_3}{M_1} + \frac{9}{4}\frac{M_2^3}{M_1^2}\right]}{6M_2^2 + \frac{1}{3}\left[7(M_4 - 3M_2^2) + 12\frac{M_2M_3}{M_1} + \frac{6M_2^3}{M_1^2}\right]}$$

where $M_i = n^{-1}[R_j - \overline{R}_j]^i$ for i = 2, 3, 4, and $M_1 = \overline{R}_j$.

Equation (7) defines the maximum likelihood estimate of α as a function of the first four moments. The value of functional parameter is

$$\lambda = 1 - \alpha$$

The empirical results listed here are limited to the cross-sectional risk return regressions pioneered by Lintner (1965) and Douglas (1969).[2] Three types of transformations were performed: (1) transformation of both independent and dependent variables; (2) transformation of dependent variable only; and (3) transformation of independent variable only. Table 3 only reports the results from the regression of the first type of transformation. After transformation most of the security returns pass the skewness test. Table 3 indicates that the coefficient of residual variance is significant in only one case in period 1 when average security return is regressed against systematic and nonsystematic risks. Further, the variance of residuals remain positive and significant only in period 3 where it is the only explanatory variable. The significance of residual variance in period 1 is primarily due to the existence of multicollinearity. The effect of multicollinearity can be seen more clearly when the average security return is regressed against the residual variance alone. The coefficient of residual variance has negative sign and is insignificant. The relationship between the coefficients in the first and third regressions in the first period can be described as follows

$$c = c^* + b^*d$$

where b^* and c^* are the coefficients of beta and residual variance in the first regression, c is the coefficient in the third regression, and d is

Table 3

Transformation of independent and dependent variables

Period 1

$$\overline{Y}_j = .00340 + .00428\beta_j - .31708\sigma^2_{ej} \qquad R^2 = .040$$
$$(3.763) \quad (4.396) \quad (-2.303)$$

$$\overline{Y}_j = .00332 + .00322\beta_j \qquad R^2 = .029$$
$$(3.667) \quad (3.737)$$

$$\overline{Y}_j = .00663 - .03258\sigma^2_{ej} \qquad R^2 = .000$$
$$(12.377) \quad (-.263)$$

Period 2

$$\overline{Y}_j = -.00066 + .00269\beta_j - .11808\sigma^2_{ej} \qquad R^2 = .012$$
$$(-.619) \quad (2.293) \quad (-.883)$$

$$\overline{Y}_j = -.00053 + .00203\beta_j \qquad R^2 = .010$$
$$(-.498) \quad (2.238)$$

$$\overline{Y}_j = .00133 + .07539\sigma^2_{ej} \qquad R^2 = .001$$
$$(2.091) \quad (.724)$$

Period 3

$$\overline{Y}_j = -.00220 + .00121\beta_j + .20626\sigma^2_{ej} \qquad R^2 = .016$$
$$(-1.614) \quad (.832) \quad (1.810)$$

$$\overline{Y}_j = -.00246 + .00263\beta_j \qquad R^2 = .009$$
$$(-1.812) \quad (2.137)$$

$$\overline{Y}_j = -.00123 + .25729\sigma^2_{ej} \qquad R^2 = .015$$
$$(-1.735) \quad (2.680)$$

the auxiliary regression coefficient of regressing beta against residual variance. The value of d is equal to $\gamma_{\beta,S^2} \frac{\sigma_\beta}{\sigma_{s^2}}$ where γ_{β,s^2} is the simple correlation between beta and residual variance, and σ_β and σ_{s^2} are the standard deviations of beta and residual variance respectively. Using the values of these elements, we found the value of d is 66.406, with $\gamma_{\beta,\sigma^2} = .47$. Thus $c = -.31708 \times (.00428 * 66.406) \cong -.033$, which is close to the coefficient of residual variance in the third regression.

B. CES Function Approximation of CAPM

The risk return relationship based on a homogeneous two parameter preference structure and an equilibrium market can be written as

$$(8) \qquad E({}_H R_j) = (1-\beta_j^*)\,{}_H R_f + \beta_j^* E({}_H R_m)$$

where ${}_H R_j$, ${}_H R_f$ and ${}_H R_m$ are rates of return in terms of true investment horizon, H. If the true investment horizon is unknown than the following relationships hold:

$$E(R_j) = [E({}_N R_j)]^\lambda$$

$$E(R_m) = [E({}_N R_m)]^\lambda$$

$$E(R_f) = [E({}_N R_f)]^\lambda$$

where ${}_N R_i$ represents the rate of returns in terms of observed horizon, and $\lambda = \frac{H}{N}$. Equation (8) becomes[3]

$$(9) \qquad E(R_j)^\lambda = (1-\beta_j^*) R_f^{\ \lambda} + \beta_j^* E(R_m)^\lambda$$

which is a CES type function of CAPM derived by Lee (1976). Equation (9) can be approximated by equation (10):

$$(10) \qquad \ln E(R_j) = (1-\beta_j^*) \ln R_f + \beta_j^* \ln E(R_m) + \frac{1}{2} \lambda \beta_j^*(1-\beta_j^*) [\ln E(R_m) - \ln R_f]^2$$

where λ measures the ratio of true horizon to observed horizon and the degree of skewness and kurtosis.[4] Equation (10) is employed to test the functional form of risk return relationship, and to test the significance of skewness and kurtosis. If the coefficient for the quadratic market excess rate of return is significant, then the higher moments, especially skewness will affect the risk return tradeoff, and the linear model is subject to specification bias.

Define

$$(11) \qquad r_j = \frac{1}{2}\lambda\beta_j^*(1-\beta_j^*)$$

The relationship between β_j and r_j can be written as

$$(12) \qquad \beta_j = \beta_j^* + br_j$$

where b is the auxiliary regression coefficient of regressing $(\ln E(R_m) - \ln(R_f))^2$ against $(\ln E(R_m) - \ln(R_f))$ and β_j is the systematic risk estimated by the linear characteristic line. β_j will be biased downward or upward depending upon the sign of b and γ.

2. Quadratic Risk Premium Market Model

Following Kraus and Litzenberger (1976), quadratic characteristic lines consistent with the three moment capital asset pricing model can be expressed as

$$(13) \qquad R_{it} - R_{ft} = c_{0i} + c_{1i}(R_{mt}-R_{ft}) + c_{2i}(R_{mt}-\bar{R}_{mt})^2 + \eta_i$$

The systematic risk β and systematic skewness γ for the ith security can be written as

$$(14) \qquad \beta_i = c_{1i} + c_{2i}\ (M_m^3/\sigma_m^2)$$

$$(15) \qquad \gamma_i = c_{1i} + c_{2i}\ [M_m^4 - \sigma_m^4]/M_m^3$$

where M_m^3 and M_m^4 are skewness and kurtosis of the market rate of return. Thus, if equation (13) is used to estimate risk return relationship, the value of systematic risk and skewness will be affected by the third and fourth moments, unless the value of c_{2i} is equal to zero. When this occurs, the linear characteristic line holds for the ith security, and the three moment capital asset pricing model reduces to the traditional two moment model.

Note that equation (13) is just one alternative to estimate the market model. The quadratic characteristic lines can be formulated in a form which is consistent with the generalized functional specification of CAPM as indicated in equation (10). This type of characteristic lines can be written as

$$(16) \qquad R_{it} - R_{ft} = c_{0i}^* + c_{1i}^*(R_{mt}-R_{ft}) + c_{2i}^*(R_{mt}-R_{ft})^2 + \tau_i$$

Again the relationships between β, γ, c_1, and c_2 are

$$(17) \qquad \beta_i = c_{1i}^* + c_{2i}^*[M_m^3/\sigma_m^2 + 2(\overline{R}_m-R_f)]$$

$$(18) \qquad \gamma_i = c_{1i}^* + c_{2i}^*[(M_m^3-\sigma_m^4)/M_m^3 + 2(\overline{R}_m-R_f)]$$

Equation (16) provides a linkage between three moment capital asset pricing model and the generalized functional transformation. The only difference between equation (16) and equation (10) is that the former specifies the risk and discrete rate of return tradeoff in a finite horizon, while the latter represents the relationship between risk and continuous compounding rate of return. Therefore, equation (16) not only is consistent with the three moment capital asset pricing model but also implicitly eliminates the statistical problems by normalizing the disturbance term to satisfy the classical assumptions for least squares.

Table 4 summarizes the regression results for equations (10), (13), and (16). c_1 and c_2 correspond to β_j^* and γ_j in equation (10). The average systematic risk estimated from the linear characteristic line is also listed for comparison. The results of three quadratic characteristic lines appear similar. The number of securities with t values greater than one for the coefficients of the quadratic term for each regression model is reported in the last column of Table 3. For those securities with the coefficients of the quadratic term significantly different from zero, the nonlinear characteristic line holds and the three moment capital asset pricing model becomes relevant.

V. Estimation Error and the Performance of Mean-Absolute-Deviation Versus Least Squares Estimators

Characteristic lines are often fitted to historical returns data by least squares. For a certain class of distributions with fat tails, the Gauss-Markov theorem no longer applies to least squares. As a result of giving more weights to outliers, least squares (LS) becomes

Table 4

Average Values of Coefficients and R^2

Period 1	$\bar{c}^1$	$\bar{c}^2$	$\bar{R}^2$	n*
Equation 10	1.0246	.9535	.25	117
Equation 13	1.0550	1.1590	.25	127
Equation 16	1.0438	1.2728	.25	128
Linear model	1.0067			
Period 2				
Equation 10	1.1008	2.1692	.28	166
Equation 13	1.1018	2.653	.29	162
Equation 16	1.0962	2.700	.29	160
Linear model	1.1036			
Period 3				
Equation 10	1.0460	.3204	.33	179
Equation 13	1.0463	.2776	.33	190
Equation 16	1.0413	.2604	.33	188
Linear model	1.0607			

*n represents the number of securities which have t values greater than one for the coefficients of the quadratic term.

extremely sample dependent. An alternative regression model less sensitive to outliers becomes more desirable. Because of its simplicity, mean-absolute-deviation (MAD) estimator has often been selected as a good candidate, at least for the preliminary estimates. By giving relatively small weights to outliers, MAD regression may provide a more efficient estimator.

To state the above argument more formally, let

(19) $$\hat{\beta}^{M}_{jt} = \beta_{jt} + \varepsilon_{jt}$$

(20) $$\hat{\beta}^{L}_{jt} = \beta_{jt} + \eta_{jt}$$

where

$\hat{\beta}^{M}_{jt}$ = MAD beta estimate for security j in period t.

$\hat{\beta}^{L}_{jt}$ = LS beta estimate for security j in period t.

β_{jt} = the assumed true value of beta risk for security j.

When the distribution of security returns exhibits high kurtosis, as evidenced by our test, it is expected that the dispersion of MAD estimates to be smaller than that of LS estimates. That is, assuming that β_{jt} and disturbance terms are uncorrelated.

(21) $$\mathrm{Var}(\hat{\beta}^{M}_{jt}) = \mathrm{Var}(\beta_{jt}) + \sigma^2_{\varepsilon jt}$$

(22) $$\mathrm{Var}(\hat{\beta}^{L}_{jt}) = \mathrm{Var}(\beta_{jt}) + \sigma^2_{\eta jt}$$

and $$\mathrm{Var}(\hat{\beta}^{M}_{jt}) < \mathrm{Var}(\hat{\beta}^{L}_{jt}).$$

This implies that $\sigma^2_{\varepsilon jt} < \sigma^2_{\eta jt}$. If the estimated beta risk in the previous period is used to predict the beta risk in the current period, the forecasting model can be written as

$$\hat{\beta}_{jt} = a + b\hat{\beta}_{jt-1} + u_{jt} \tag{23}$$

where $\hat{\beta}_{jt}$ and $\hat{\beta}_{jt-1}$ are either MAD or LS estimates. Again the traditional Mincer and Zarnowitz method can be used to analyze the forecasting errors. However, the independent and dependent variables are now measured with error. This proxy error will lead to the underestimation of b coefficient. To see this more clearly, let

$$b^M = \frac{\mathrm{Cov}(\hat{\beta}^M_{jt}, \hat{\beta}^M_{jt-1})}{\mathrm{Var}(\hat{\beta}^M_{jt-1})} \tag{24}$$

$$b^L = \frac{\mathrm{Cov}(\hat{\beta}^L_{jt}, \hat{\beta}^L_{jt-1})}{\mathrm{Var}(\hat{\beta}^L_{jt-1})} \tag{25}$$

where b^M and b^L represent the slope coefficients of equation (23) for MAD and LS regression estimates respectively. Substitute equations (19), (20), (21) and (22) into equation (24) and (25) and assume that the covariances for cross-period disturbance terms are zero (no serial correlation),

$$b^M = \frac{\mathrm{Cov}(\beta_{it}, \beta_{jt-1})}{\mathrm{Var}(\beta_{it-1}) + \sigma^2_{\varepsilon jt-1}} = \frac{b}{1 + \dfrac{\sigma^2_{\varepsilon jt-1}}{\mathrm{Var}(\beta_{jt-1})}} \tag{26}$$

$$b^L = \frac{\mathrm{Cov}(\beta_{jt}, \beta_{jt-1})}{\mathrm{Var}(\beta_{it-1}) + \sigma_{\eta jt-1}} = \frac{b}{1 + \dfrac{\sigma^2_{\eta jt-1}}{\mathrm{Var}(\beta_{jt-1})}} \tag{27}$$

Since $\sigma^2_{\varepsilon jt-1} < \sigma^2_{\eta jt-1}$ and both are positive, $b > b^M > b^L$ where b represents the estimated coefficient of equation (23) when there is no proxy error.

The value of the coefficient of determination of equation (24) is also affected. Since there is only one explanatory variable, the value of the squared simple correlation is equal to that of the coefficient of determination.

$$(28) \qquad R^2 = \frac{\mathrm{Cov}(\hat{\beta}_{jt},\hat{\beta}_{jt-1})^2}{\sigma^2_{\hat{\beta}jt}\sigma^2_{\hat{\beta}it-1}}$$

For simplicity, assume that the variances of β_{jt} and disturbance terms are constant over time. Denote R^2_M and R^2_L as coefficients of determination for MAD and LS estimates respectively. Then

$$(29) \qquad R^2_M = \frac{\mathrm{Cov}(\beta_{jt},\beta_{jt-1})^2}{(\sigma^2_{\beta jt}+\sigma^2_{\varepsilon jt})^2}$$

$$(30) \qquad R^2_L = \frac{\mathrm{Cov}(\beta_{it},\beta_{jt-1})^2}{(\sigma^2_{\beta it}+\sigma^2_{\eta it})^2}$$

which implies that $R^2_M > R^2_L$. Rewrite the Mincer-Zarnowitz MSE decomposition as

$$\mathrm{MSE} = (\bar{\beta}_t-\bar{\beta}_{t-1})^2 + (1-b)\hat{\sigma}^2_{\hat{\beta}t-1} + (1-R^2)\hat{\sigma}^2_{\hat{\beta}t}$$

Thus it is expected that both inefficiency and random components to be smaller for the MAD estimates.

Two methods are often used to estimate the MAD regression coefficients. The first one is the iternative procedure proposed by Karst (1958) and Sharpe (1971). Sharpe's method is based on a normal equation expressing the sum of absolute deviation as a piecewise convex function of the beta estimate. Although Sharpe designed an algorithm to solve problems in only two dimensions, his method can be extended to multi-dimensional problems. However, as the number of independent variable increases, the computing procedures become cumbersome.

The second method is the linear programming model suggested by Charnes, Cooper and Ferguson (1955), and Wagner (1959). Linear programming is especially suitable for regression with more than one independent variable. A primal algorithm proposed by Barrodale and Roberts is adopted here. A feature of this algorithm is its ability to pass through several simplex vertices at each iteration. Our experience shows this algorithm is more efficient than the generally accepted dual algorithm.

In order to examine the performance of MAD relative to LS estimates more closely, data are once again separated into two groups, that is, returns distribution with and without significant high kurtosis. It is well known that fat tail is associated with high kurtosis. This type of grouping allow us to evaluate the performance of MAD estimates more easily. The performance of these two estimators are measured according to their forecasting capability.

The empirical results for MAD estimates are reported in Table 5. When compared with Table 2, it is somewhat surprising that LS estimates

Table 5

The Stability of Mean-Absolute-Deviation Beta Estimate*

Group	R^2	Slope	Mean Square Errors bias	inefficiency	random	total
(1) Normality						
1959-65 1966-72	.23	.516	.0079032 5%	.0334055 20%	.1267178 75%	.1680265
1966-72 1973-79	.31	.473	.0070728 5%	.0450288 34%	.0795368 61%	.1316385
(2) Significant kurtosis						
1959-65 1966-72	.22	.550	.0265136 12%	.0309154 14%	.1613588 74%	.2187879
1966-72 1973-79	.25	.490	.0000435 0%	.0582932 27%	.1599715 73%	.2183082
(3) All companies						
1959-65 1966-72	.22	.524	.012528 7%	.033164 18%	.138047 75%	.183739
1966-72 1973-79	.28	.478	.003696 2%	.048597 32%	.101566 66%	.153859

*The estimated regression model is: $Beta_t = b_0 + b_1 \; Beta_{t-1}$

perform better than MAD estimates. While the general patterns of MSE are similar, in most cases, LS estimates provide smaller forecasting error for each MSE component: bias, inefficiency or random. The R squares are also higher and the b coefficients of equation (23) are larger for each group indicating that LS estimates are more efficient than MAD estimates.

To examine the relation between MAD and LS estimates, we regress the former against the latter. This type of regression provides a scheme to see how close these two estimates are. As shown in Table 6, for the first group the difference between these two estimates is mostly random. However, for the second group, the bias portion of MSE increases substantially. This phenomenon may reflect the fact that LS and MAD are mean and median regression estimators respectively, and in the presence of high kurtosis the difference between both estimates becomes more significant. To provide further insights on the performance of these two estimators, the cross-sectional risk-return tradeoff and the relationship between performance measure and risk were also examined. Again, in every case, LS estimates provided better or at least no worse results. Our findings therefore question the superiority of MAD estimator in estimating the market model for the monthly returns.

There are at least several reasons to explain why MAD estimator did not outperform LS estimator. First, the existence of misspecification is likely to affect the regression results. As shown in Table 4, for a number of securities the coefficient representing skewness of

Table 6

Mean-Absolute-Deviation versus Least Squares Beta Estimates*

Group	R^2	Slope	Mean Square Errors			
			bias	inefficiency	random	total
(1) Normality						
1959-65	.87	1.006	.0000588 0%	.0000047 0%	.0175168 100%	.0175804
1966-72	.88	.968	.0000122 0%	.0001479 1%	.0193189 99%	.0194791
1973-79	.83	1.036	.0020884 9%	.0001211 1%	.0200317 90%	.0222412
(2) Significant kurtosis						
1959-65	.79	.937	.0081234 21%	.0005458 1%	.0306277 78%	.0392970
1966-72	.84	.916	.0066912 15%	.0015721 4%	.0349996 81%	.0432630
1973-79	.79	.957	.0066912 17%	.0002352 1%	.0314874 82%	.0384139
(3) All companies						
1959-65	.84	.977	.001176 5%	.000064 0%	.023185 95%	.024425
1966-72	.86	.944	.000338 2%	.000533 2%	.024709 96%	.025580
1973-79	.78	.964	.001317 4%	.000151 0%	.031170 96%	.032638

*The estimated regression model is: $Beta_t(MAD) = a_0 + a_1\, Beta_t(LS)$

the returns distribution is significant. Second, when the distribution of stock returns is skew, the linear model no longer represents the appropriate functional form. This being the case, it is difficult to expect the MAD estimator to improve the empirical estimation. Third, the error in variable problem resulted from market index proxy may considerably limit the usefulness of the MAD estimator. Finally, Zellner (1976) shows that for the multivariate t distribution, LS estimator is the maximum likelihood estimator. Blattarg and Gonedes (1974) have provided evidence that the t distribution fits the stock returns distribution very well. The sample variances of security returns examined in the study are all well behaved and have finite value. Our results tend to reject the hypothesis of stable class distribution.

To further examine the performance of the MAD estimator, a group was formed to consist the data with high kurtosis and no skewness. Again, the LS estimator outperforms the MAD estimator in every aspect. Therefore, this study provides further evidence that even in the existence of kurtosis, the MAD estimator is not necessarily efficient. In their Monte Carlo study, Ruppore and Carroll (1980) also indicate that MAD estimator will outperform LS estimator only under the restriction that kurtosis is really high.

V. Summary

In accordance with the modern statistical theory and econometric methods, this paper investigates the impacts of skewness and kurtosis on risk estimation and determination. First, impacts of kurtosis on the total risk estimation are analytically and empirically studied. Second, the relationships between stability of beta coefficients and the third

and fourth moments of security rate of return are analyzed in detail. Third, the impacts of skewness and kurtosis on risk decomposition are determined. Finally, the concepts and measurement of estimation errors are used to determine whether the OLS or the MAD method should be used to estimate beta coefficients. In sum, this paper has demonstrated that both skewness and kurtosis of security rates of return should be concerned in estimating and determining the risk in capital asset pricing.

Footnotes

[1]Scott and Horvath (1980) shows that investor exhibiting consistent risk aversion, strict consistency of moment perference and positive perference for positive skewness will have negative perference for kurtosis.

[2]Our model differs from that of Friend and Westerfield (1980,81). The problem of market index is not the major concern of this study, while Friend and Westerfield did not consider the effect of kurtosis.

[3]The N used to indicate the observe horizon is omitted to simplify the notation.

[4]Cootner (1962) and others show that the degree of skewness and kurtosis is a function of the length of observated horizon.

References

Andrews, D. F., "A Robust Method for Multiple Linear Regression," Technometrics 16(1974) 523-31.

Blattberg, R. C. and Gonedes, "A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices," Journal of Business 47(1974) 244-280.

Blume, M. E., "The Assessment of Portfolio Performance: An Application of Portfolio Theory," Doctoral Dissertation, University of Chicago (1968).

Blume, M. E., "On the Assessment of Risk," Journal of Finance (March 1971) 1-10.

Blume, M. E., "Betas and Their Regression Tendencies," Journal of Finance (March 1975) 785-795.

Box, G. E. P. and Cox, D. R., "An Analysis of Transformations," Journal of the Royal Statistical Society, Series B, 26(1964) 211-243.

Charles, A., Cooper, W. W. and Ferguson, R. O., "Optimal Estimation of Executive Compensation by Linear Programming," Management Science (1955) 138-151.

Cooley, T. F. and E. C. Prescott, "Systematic (Non-Random) Variation Models Varying Parameter Regression: A Theory and Some Applications," Annals of Economics and Social Measurement 2 (February/April 1973) 463-73.

Cootner, P. H., "Stock Prices: Random vs. Systematic Changes," Industrial Management Review, Spring 1962.

Cornell, B. and Dietrich, J. K., "Mean-Absolute Deviation Versus Least-Squares Regression Estimation of Beta Coefficients," Journal of Financial and Quantitative Analysis, 13(1978) 123-131.

Douglas, G. W., "Risk in the Equity Markets: An Empirical Appraisal of Market Efficiency," Yale Economic Essays 9 (Spring 1969) 3-45.

Fabozzi, F. J., Francis J. C. and Lee, C. F. (1980), "Generalized Form for Mutual Fund Returns," Journal of Financial and Quantitative Analysis, 15, pp. 1107-1120.

Fama, E. F., "The Behavior of Stock Market Prices," Journal of Business (January 1965) 34-105.

Fama, E. F., Foundations of Finance, New York 1976.

Fielitz, B. D., "Stationarity of Random Data: Some Implications for the Distribution of Stock Price Changes," Journal of Financial and Quantitative Analysis 6 (June 1971) 1025-34.

Fisher, R. A., Statistical Methods for Research Workers, Oliver and Boyd, Edinburgh, fourth edition, 1932.

Francis, J. C., "Statistical Analysis of Risk Measures for NYSE Stocks," Journal of Financial and Quantitative Analysis 14 (December 1979) 981-997.

Friend, I. and Blume, M., "Measurement of Portfolio Performance under Uncertainty," The American Economic Review 60 (1970) 561-75.

Friend, I. and Westerfield, R., "Risk and Capital Asset Pricing," Journal of Banking and Finance 5(1981) 291-315.

Friend, I. and Westerfield, R., "Co-Skewness and Capital Asset Pricing," The Journal of Finance 25 (September 1980) 897-913.

Galai, D. and Masulis, R. W., "The Option Pricing Model and the Risk Factor of Common Stock," Journal of Financial Economics (March 1976) 53-81.

Hamada, R. S., "The Effects of the Firm's Capital Structure on the Systematic Risk of Common Stocks," Journal of Finance (May 1972) 435-452.

Haitovsky, Y., "A Note on the Maximization of R Square," The American Statistician (February 1969) 20-21.

Hsu, D. A., "Stochastic Instability and the Behavior of Stock Prices," Unpublished Ph.D. dissertation, University of Wisconsin, Madison, 1973.

Hsu, D., Miller, R. B. and Wichern, D. W., "On the Stable Paretian Behavior of Stock-Market Prices," Journal of the American Statistical Association 69 (March 1974) 108-113.

Huber, P. J., "Robust Regression, Asyptotics, Conjectures and Monte Carlo," Annals of Statistics 1 (1973) 799-821.

Karst, Otto J., "Linear Curve Fitting Using Least Deviations," Journal of the American Statistical Association (March 1958) 118-132.

Koenker, R. and Bassett, G. J., "Regression Quantiles," Econometrica 46 (January 1978) 33-50.

Klemkosky, R. C. and Martin, J. D., "The Adjustment of Beta Forecasts," The Journal of Finance 30 (1975) 1123-1128.

Kraus, A. and Litzenberger, R. H., "Skewness Preference and the Evaluation of Risk Assets," Journal of Finance, September 1976.

Jensen, M. C., "Risk, the Pricing of Capital Assets, and the Evaluation of Investment Portfolio," Journal of Business (April 1969) 167-247.

Levy, R. A., "On the Short-Term Stationarity of Beta Coefficients," Financial Analysis Journal (November/December 1971) 55-62.

Lee, C. F., "Investment Horizon and the Functional Form of the Capital Asset Pricing Model," The Review of Economics and Statistics, August 1976.

Lintner, J., "Security Prices, Risk and Maximal Gains from Diversification," Journal of Finance 20 (1965) 587-616.

Mandelbrot, B., "The Variation of Certain Speculative Prices," Journal of Business 36 (October 1963) 394-419.

Mandelbrot, B., "The Variation of Some Other Speculative Prices," Journal of Business 40 (October 1967) 393-413.

Meyer, J. R. and Glauber, R. R., Investment Decisions, Economic Forecasting, and Public Policy, Harvard Business School Press, Cambridge, Massachusetts (1964).

Miller, R. B., Wichern, D. W. and Hsu, D. A., "Changes of Variance in Time Series Models," Technical Report No. 284, Department of Statistics, University of Wisconsin, Madison, 1971.

Miller, M. H. and Scholes, M., "Rate of Return in Relation to Risk: A Reexamination of Some Recent Findings," in M. C. Jensen, ed., Studies in the Theory of Capital Market, New York: Praeger 1972.

Officer, R. R., "The Distribution of Stock Returns," Journal of the American Statistical Association 67 (December 1972) 807-12.

Mincer, J. and Zarnowitz, V., "The Evaluation of Economic Forecasts," Economic Forecasts and Expectation: Analysis of Forecasting Behavior and Performance, edited by J. Mincer, New York: National Bureau of Economic Research (1969).

Ruppert, D. and Carroll, R. J., "Trimmed Least Squares Estimation in the Linear Model," Journal of the American Statistical Association (December 1980) 828-838.

Sargent, T. J., "Regressions in the Context of Stable Disturbances," National Bureau of Economic Research Forty-Ninth Annual Report (October 1969) 107-108.

Schiller, R. J., "Do Stock Prices Move Too High to be Justified by Subsequent Changes in Dividends?," The American Economic Review 71 (June 1981) 421-436.

Scott, R. C. and Horvath, P. A., "On the Direction of Preference for Moments of Higher Order than the Variance," The Journal of Finance 25 (September 1980) 911-19.

Sharpe, W. F., "Mean-Absolute-Deviation Characteristic Lines for Securities and Portfolios," Management Science (October 1971) B1-B13.

Snedecor, G. W. and Cochran, W. G., Statistical Methods, The Iowa State University Press, (1967) 84-89.

Taylor, L. D., "Estimation by Minimizing the Sum of Absolute Errors," Frontiers in Econometric, edited by P. Zarembka, New York: Academic Press (1973).

Wiginton, J. C., "Portfolio Analysis under Stable Paretian Hypothesis: An Empirical Evaluation," New York University Capital Markets Research Program Working Paper (1968).

Wise, J., "Linear Estimators for Linear Regression Systems Having Infinite Residual Variance," Berkeley-Stanford Mathematical Economics Seminar (1963).

Zellner, A., "Bayesian and Non-Bayesian Analysis of the Regression Model with Multivariate Student t Error Terms," Journal of the American Statistical Association, 71, 400-405.